

# Monitoring COVID-related Social Disruption using Semantic Knowledge Extraction

## OVERVIEW

In emergencies and complex environments, *everything* is a priority, but policymakers must choose what to prioritize given finite resources for effective response. Information overload can be even more paralyzing than information scarcity. Artificial Intelligence (AI) can help us to better process large amounts of information and identify pandemic relevant events from massive feeds of unstructured data. The Cline Center for Advanced Social Research will apply semantic knowledge extraction methods to two extreme-scale global news sources (one proprietary, the other provided by an external vendor) to identify pandemic-related social disruptions, socio-economic effects, and policy responses. This effort will be supported by collaboration with experts across the GReVD consortium. The resulting data will be published for policy and research applications.



## ABSTRACT

The relationship between a pandemic and social order/disorder is complex. The spread of a contagion may exacerbate conflicts between social groups, between state and society, or even across national borders. If handled well, a pandemic response might contribute to social cohesion and improve state-society relationships. If handled poorly, pandemics can have massive, multi-faceted, and enduring negative effects. These impacts may not be immediately apparent—but with large-scale mortality shocks, eventually, they will be substantial. Grievances may lay dormant until the pandemic has passed, and then explode into open conflict after a seemingly-trivial triggering event. Even in the best case, negative economic and development effects will persist for years, and *how* governments respond to the pandemic may be assessed and reassessed by citizens over the coming decade. In the most vulnerable, fragile nations, the impact of COVID on the Sustainable Development Agenda could easily persist until 2030.

Just as managing the health effects of a pandemic requires understanding, mitigating, and eventually containing the disease, mitigating the pandemic's societal effects requires discovering, documenting, understanding risk and resilience factors, as well as policy interventions that effectively reduce pandemic-related economic disruptions, political instability, and conflict processes.

This initiative will scale up and enhance algorithmic knowledge extraction methods already being developed at the Cline Center. Current-generation machine-generated event data are ill-suited to pandemic-related analyses. They are also based on outdated technologies—and it took nearly a decade to move them from the lab to unclassified policy applications. We propose a software development 'sprint' that would rapidly transition current-generation AI technologies from the workbench to open-access production.

The project will produce a variety of data resources, algorithms, and tools in a risk-minimizing spiral development process. Within two months of project start, we will release data based on incremental improvements to existing event data generation tools. Within three months of project start, deep learning-based knowledge extraction tools will be scaled up and deployed. Publication of data created by these systems will continue to the end of the funding period. Improved AI systems and benchmarking data will be publicly released by the end of the funding period.

## OUTPUTS

*Geo-located Event Records (Daily)*  
*Social Graph of entities—governments, companies, individuals, etc.—and relationships between them*  
*Public Data, released to researchers*  
*Resources, algorithms and tools for analysis and monitoring of pandemic responses*

Envisioned outputs include:

- Daily pandemic-relevant event records for every city, town, or county-equivalent geographical entity mentioned in news feeds from around the world. The system will capture both large and small-scale events on a global scale.
- Data on institutional and community actions will be publicly released rapidly—in time for analysis by research teams capable of rapidly assessing community resilience and vulnerability, prototyping and assessing policy interventions designed to curb the destabilizing societal impacts of COVID-19.
- We will also publish data on the entities and relationships mentioned in pandemic-related news—essentially, a massive graph of the social networks engaged in pandemic response
- The initiative will make available data, algorithms, and tools useful for monitoring and predicting pandemic-related needs impacts while the pandemic is still underway. These resources will directly support public-sector entities tasked with monitoring and mitigating the social impacts of the disease, and enable other reserachers to develop innovative software and analytical methods.

These development efforts are directly relevant to future GReVD objectives as identified in the GReVD Challenges Framework. Many GReVD partners—including the Cline Center—have have noted the inadequacies of current machine approaches to accurately extracting complex event data that describe conflicts, pandemics and violent deaths generally (see the [Cline Center technical paper](#) on using machine methods to document conflict with Boko Haram in West Africa).

Assessing algorithmic approaches and how close AI can come to human judgement requires integrating the wisdom of expert human analysts into the learning process, and collaboration across social science and engineering disciplines. GReVD is uniquely situated to bring human coders and subject matter experts together with machine learning to determine the best ways to deploy machines in order to supplement or, eventually, replicate human judgement (see [Cline Center/Global Terrorism Database \(GTD\)/Small Arms Survey \(SAS\) study](#) on machine-in-the-loop challenges and opportunities).

The ability to detect and accurately record basic information about policy changes and small-scale contentious, violent, or repressive actions will allow researchers, governments, and citizens to rapidly assess the effects of policies at the level of cities, states or provinces, and entire countries. The raw data and tools we propose to develop will support ongoing transnational efforts to craft, mobilize, and test the effectiveness of policy interventions to minimize the societal impact of the COVID-19 pandemic.